ELSEVIER

Linear Algebra and its Applications 337 (2001) 1–20

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Banded matrices and difference equations

Werner Kratz*

Abteilung Angewandte Analysis, Universität Ulm, D-89069 Ulm, Germany

Received 6 December 2000; accepted 16 March 2001

Submitted by R.A. Brualdi

Abstract

In this paper we consider *discrete* Sturm–Liouville eigenvalue problems of the form

$$L(y)_k := \sum_{\mu=0}^{n} (-\Delta)^{\mu} \left\{ r_{\mu}(k) \Delta^{\mu} y_{k+1-\mu} \right\} = \lambda y_{k+1}$$

for $0 \leqslant k \leqslant N-n$ with $y_{1-n} = \cdots = y_0 = y_{N+2-n} = \cdots = y_{N+1} = 0$, where N and n are integers with $1 \leqslant n \leqslant N$ and under the assumption that $r_n(k) \neq 0$ for all k. These problems correspond to eigenvalue problems for symmetric, banded matrices $\mathscr{A} \in \mathbb{R}^{(N+1-n)\times(N+1-n)}$ with bandwidth $2n+1$. We present the following results:

1. an inversion formula, which shows that *every* symmetric, banded matrix corresponds uniquely to a Sturm–Liouville eigenvalue problem of the above form;
2. a formula for the characteristic polynomial of $\mathscr{A}$, which yields a *recursion* for its calculation;
3. an *oscillation theorem*, which generalizes well-known results on tridiagonal matrices.

These new results can be used to treat numerically the algebraic eigenvalue problem for symmetric, banded matrices without reduction to tridiagonal form. © 2001 Elsevier Science Inc. All rights reserved.

AMS classification: 39A10; 39A12; 65F15; 15A18

Keywords: Sturm–Liouville equations; Banded matrices; Sturmian chains; Symmetric eigenvalue problem; Poincaré separation theorem; Discrete oscillation; Hamiltonian difference systems; Jacobi's formula; Picone's identity

* Fax: +49-731-502-3619.

E-mail address: kratz@mathematik.uni-ulm.de, URL: http://www.mathematik.uni-ulm.de/m5/persons/kratz.html (W. Kratz).

PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 3 2 8 - 7

1. Introduction

We consider *discrete* Sturm–Liouville eigenvalue problems (with eigenvalue parameter λ) of the form

$$L(y)_k := \sum_{\mu=0}^{n} (-\Delta)^{\mu} \left\{ r_{\mu}(k) \Delta^{\mu} y_{k+1-\mu} \right\} = \lambda y_{k+1}$$

for $0 \leqslant k \leqslant N-n$, where $\Delta y_k = y_{k+1} - y_k$, and with the boundary conditions

$$y_{1-n} = \cdots = y_0 = y_{N+2-n} = \cdots = y_{N+1} = 0,$$

where N and n are fixed integers with $1 \leqslant n \leqslant N$ and where we assume that

$$r_n(k) \neq 0 \quad \text{for all } k.$$

These problems correspond to eigenvalue problems for symmetric, banded matrices $\mathscr{A}$ of size $(N+1-n) \times (N+1-n)$ with bandwidth $2n+1$. In particular, $\mathscr{A}$ is *tridiagonal* in the case $n = 1$.

In this paper we derive the following results:

- An *inversion formula* (Theorem 1). This identity can be used to calculate the matrix $\mathscr{A}$ when the above discrete Sturm–Liouville operator is given and vice versa. Hence, *every* symmetric, banded matrix with bandwidth $2n+1$ corresponds uniquely to such a (discrete and self-adjoint) Sturm–Liouville operator.
- A formula for the characteristic polynomial of $\mathscr{A}$ (Theorem 2). This result yields also a *recursion* for its calculation. In the case $n = 1$ we obtain the well-known algorithm, which is commonly used in numerical analysis to handle eigenvalue problems for tridiagonal matrices (cf. [9, p. 305]; [13, p. 134]; [14, p. 299]).
- An *oscillation theorem* (Theorem 3). This result generalizes corresponding well-known results for tridiagonal matrices (cf. [9, Theorem 8.5-1]; [13, Satz 4.9]; or [14, pp. 300–301]).

Our method (for example the use of *Picone's identity* (Lemma 5) and most of our results have continuous counterparts along the lines of the book [11] (cf. also [12]). Moreover, our results may be applied to every ("full") real and symmetric matrix (compare Remark 1 (ii) below).

The above-mentioned results are known only in the case $n = 1$, i.e., where the matrix $\mathscr{A}$ is tridiagonal. In the general case these results are new, and they can be used (similarly as in the case $n = 1$, cf. [14, Chapter 5 pp. 299, 315]; compare also the comments in Section 6 after Remark 4) to treat numerically the algebraic eigenvalue problem for symmetric, banded matrices (without reduction to tridiagonal form via Givens' or Householder's method as described in [9, Section 8.2], or [14, Chapter 8]; for general aspects concerning banded matrices see also [1]). But, of course, the *numerical aspects* of our theory have to be analyzed in detail and compared with the existing methods. This will be a separate project.

2. Banded matrices and Sturm–Liouville difference operators

Let $n \in \mathbb{N}$, and let be given reals $r_\mu(k)$ for $\mu \in \{0, 1, \dots, n\}$ and $k \in \mathbb{Z}$. Then, for $y = (y_k)_{k\in\mathbb{Z}}$, we consider the Sturm–Liouville difference operator $L(y)$ defined by

$$L(y)_k := \sum_{\mu=0}^{n} (-\Delta)^\mu \left\{ r_\mu(k) \Delta^\mu y_{k+1-\mu} \right\} \quad \text{for } k \in \mathbb{Z}, \tag{1}$$

where Δ is the forward difference operator, i.e., $\Delta w_k := w_{k+1} - w_k$, which will always operate with respect to the variable k. Of course, the operator L is given by a multiplication of the sequence y by an infinite, banded matrix $\mathscr{A}$ with bandwidth $2n + 1$. The special (self-adjoint) form of the operator implies that the matrix $\mathscr{A}$ is symmetric. Moreover, $\mathscr{A}$ is invertible, i.e., the $r_\mu(k)$ can be expressed in terms of the matrix elements of $\mathscr{A}$. Hence, there is a unique correspondence between Sturm–Liouville operators L, given by (1), and symmetric, banded matrices $\mathscr{A}$ with bandwidth $2n + 1$. This statement, including an explicit formula for the inverse of $\mathscr{A}$, is the contents of our first theorem.

Theorem 1 (Inversion). *Let be given an operator L by formula* (1). *Then the following holds*:

(i)

$$L(y)_k = (\mathscr{A} y)_{k+1} \quad \textit{for } k \in \mathbb{Z},$$

where $\mathscr{A} = (a_{\mu\nu})$ is a symmetric, banded matrix with bandwidth $2n + 1$, given by

$$\begin{aligned} a_{k+1,k+1+t} &= (-1)^t \sum_{\mu=t}^{n} \sum_{\nu=t}^{\mu} \binom{\mu}{\nu} \binom{\mu}{\nu - t} r_\mu(k + \nu), \\ a_{k+1,k+1-t} &= (-1)^t \sum_{\mu=t}^{n} \sum_{\nu=0}^{\mu-t} \binom{\mu}{\nu} \binom{\mu}{\nu + t} r_\mu(k + \nu), \end{aligned} \tag{2}$$

for $0 \leqslant t \leqslant n$ and all $k \in \mathbb{Z}$.

(ii)

$$\begin{aligned} r_\mu(k + \mu) = (-1)^\mu \sum_{s=\mu}^{n} \Bigg\{ \binom{s}{\mu} a_{k+1,k+1+s} + \sum_{l=1}^{s-\mu} \frac{s}{l} \binom{\mu + l - 1}{l - 1} \\ \times \binom{s - l - 1}{s - \mu - l} a_{k+1-l,k+1-l+s} \Bigg\}, \end{aligned} \tag{3}$$

for $0 \leqslant \mu \leqslant n$ and all $k \in \mathbb{Z}$.

Proof. First, using the formula for finite differences, i.e.,

$$\Delta^{\mu} w_k = \sum_{\nu=0}^{\mu} \binom{\mu}{\nu} (-1)^{\mu-\nu} w_{k+\nu}, \tag{4}$$

we obtain from (1) that

$$\begin{aligned}
L(y)_k &= \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \binom{\mu}{\nu} (-1)^{\nu} r_{\mu}(k+\nu) \Delta^{\mu} y_{k+1-\mu+\nu} \\
&= \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \sum_{s=0}^{\mu} \binom{\mu}{\nu} \binom{\mu}{s} (-1)^{\nu+\mu-s} r_{\mu}(k+\nu) y_{k+1-\mu+\nu+s} \\
&= \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \sum_{s=0}^{\mu} \binom{\mu}{\nu} \binom{\mu}{s} (-1)^{\nu+s} r_{\mu}(k+\nu) y_{k+1+\nu-s} \\
&\quad \text{(replacing } s \text{ by } \mu - s) \\
&= \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \sum_{t=0}^{\nu} \binom{\mu}{\nu} \binom{\mu}{\nu-t} (-1)^{t} r_{\mu}(k+\nu) y_{k+1+t} \\
&\quad + \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \sum_{t=1}^{\mu-\nu} \binom{\mu}{\nu} \binom{\mu}{\nu+t} (-1)^{t} r_{\mu}(k+\nu) y_{k+1-t} \\
&\quad \text{(putting } t = \nu - s \text{ for } s \leqslant \nu \text{ and } -t = \nu - s \text{ for } s > \nu) \\
&= \sum_{t=0}^{n} \left\{ (-1)^{t} \sum_{\mu=t}^{n} \sum_{\nu=t}^{\mu} \binom{\mu}{\nu} \binom{\mu}{\nu-t} r_{\mu}(k+\nu) \right\} y_{k+1+t} \\
&\quad + \sum_{t=1}^{n} \left\{ (-1)^{t} \sum_{\mu=t}^{n} \sum_{\nu=0}^{\mu-t} \binom{\mu}{\nu} \binom{\mu}{\nu+t} r_{\mu}(k+\nu) \right\} y_{k+1-t},
\end{aligned}$$

and this last equation implies formula (2). Moreover, $\mathscr{A}$ is symmetric, because, by (2),

$$\begin{aligned}
a_{k+t-1,k+1} &= (-1)^{t} \sum_{\mu=t}^{n} \sum_{\nu=t}^{\mu} \binom{\mu}{\nu} \binom{\mu}{\nu-t} r_{\mu}(k-t+\nu) \\
&\quad \text{(putting } s = \nu - t) \\
&= (-1)^{t} \sum_{\mu=t}^{n} \sum_{s=0}^{\mu-t} \binom{\mu}{s+t} \binom{\mu}{s} r_{\mu}(k+s) \\
&= a_{k+1,k+1-t}.
\end{aligned}$$

For the proof of assertion (ii) we require the following lemma.

Lemma 1. *The polynomials* $f_\mu^s(x)$, *defined by*

$$f_\mu^s(x) := (-1)^{s-\mu}\left\{\binom{s}{\mu} + \sum_{l=1}^{s-\mu}\frac{s}{l}\binom{\mu+l-1}{l-1}\binom{s-l-1}{s-\mu-l}x^l\right\} \tag{5}$$

for $0 \leqslant \mu \leqslant s$, *satisfy the relation*

$$\sum_{\mu=t}^{s}\sum_{\nu=0}^{\mu-t}\binom{\mu}{\nu}\binom{\mu}{\mu-t-\nu}x^\nu f_\mu^s(x) = \delta_{st} \tag{6}$$

for $0 \leqslant t \leqslant s$, *where* δ_{st} *denotes the Kronecker-symbol.*

Proof of Lemma 1. Let $0 \leqslant t \leqslant s$, and define a polynomial $f(x)$ of degree at most $s-t$ by

$$f(x) := \sum_{\mu=t}^{s}\sum_{\nu=0}^{\mu-t}\binom{\mu}{\nu}\binom{\mu}{\mu-t-\nu}x^\nu f_\mu^s(x) = \sum_{k=0}^{s-t}\alpha_k x^k,$$

where the $f_\mu^s(x)$ are given by (5). We have to prove that

$$f(x) = \delta_{st},$$

i.e.,

$$\alpha_0 = \delta_{st} \quad \text{and} \quad \alpha_k = 0 \text{ for } 1 \leqslant k \leqslant s-t.$$

This will be shown in the following by using (5), elementary identities for binomial coefficients (which are direct consequences of their definition) and by comparing coefficients of the power series of suitable binomial series.

(i)

$$\begin{aligned}\alpha_0 &= \sum_{\mu=t}^{s}\binom{\mu}{0}\binom{\mu}{\mu-t}(-1)^{s-\mu}\binom{s}{\mu}\\ &= \sum_{\nu=0}^{s-t}(-1)^\nu\binom{s-\nu}{t}\binom{s}{s-\nu}\\ &= \binom{s}{t}\sum_{\nu=0}^{s-t}(-1)^\nu\binom{s-t}{\nu}\\ &= \binom{s}{t}\delta_{st}\\ &= \delta_{st}.\end{aligned}$$

(ii)

$$
\begin{aligned}
\alpha_{s-t} &= \sum_{\mu=t}^{s-1} \binom{\mu}{\mu-t}\binom{\mu}{0}(-1)^{s-\mu}\frac{s}{s-\mu}\binom{s-1}{s-\mu-1}\binom{\mu-1}{0} \\
&\quad + \binom{s}{s-t}\binom{s}{0}(-1)^{s-s}\binom{s}{s} \\
&= \sum_{\mu=t}^{s} \binom{\mu}{\mu-t}(-1)^{s-\mu}\binom{s}{\mu} \\
&= \delta_{st} \\
&= 0 \quad \text{by (i) if } s-t \geqslant 1.
\end{aligned}
$$

(iii) Now, let $1 \leqslant k \leqslant s-t-1$. Then

$$
\begin{aligned}
\alpha_k = {} & \sum_{\mu=t+k}^{s} \binom{\mu}{k}\binom{\mu}{\mu-t-k}(-1)^{s-\mu}\binom{s}{\mu} \\
& + \sum_{r=0}^{k-1}\sum_{\mu=t+r}^{s-k+r} \binom{\mu}{r}\binom{\mu}{\mu-t-r}(-1)^{s-\mu} \\
& \times \frac{s}{k-r}\binom{\mu+k-r-1}{k-r-1}\binom{s-k+r-1}{s-\mu-k+r} \\
= {} & \sum_{\nu=0}^{s-t-k} \binom{\nu+t+k}{k}\binom{\nu+t+k}{\nu}(-1)^{s+t+k+\nu}\binom{s}{\nu+t+k} \\
& + \sum_{r=0}^{k-1}\sum_{\nu=0}^{s-k-t} \binom{\nu+t+r}{r}\binom{\nu+t+r}{\nu}(-1)^{s+\nu+t+r} \\
& \times \frac{s}{k-r}\binom{\nu+t+k-1}{k-r-1}\binom{s-k+r-1}{s-\nu-t-k} \\
= {} & \mathrm{I} + \mathrm{II}.
\end{aligned}
$$

$(\text{iii})_1$

$$
\begin{aligned}
\mathrm{I} &= \sum_{\nu=0}^{s-t-k} (-1)^{s+t+k+\nu}\binom{\nu+t+k}{k}\binom{s-k-t}{\nu}\binom{s}{t+k} \\
&= \binom{s}{t+k}\sum_{\mu=t}^{s-k}(-1)^{s-k-\mu}\binom{s-k-t}{s-k-\mu}\binom{\mu+k}{\mu} \\
&= \binom{s}{t+k} \times (s-k)\text{th coefficient of } (1-x)^{s-k-t}(1-x)^{-k-1}
\end{aligned}
$$

$$= (-1)^{s-k}\binom{s-2k-t-1}{s-k}\binom{s}{t+k}.$$

(iii)$_2$

$$\begin{aligned}\mathrm{II} = {} & \sum_{r=0}^{k-1}\sum_{\nu=0}^{s-k-t}(-1)^{s-k-t-\nu}\binom{s-k-t}{s-k-t-\nu}\binom{\nu+t+k-1}{\nu+t}\\ & \times(\nu+t+r)(-1)^{r+k}\binom{s-k+r}{r+t}\binom{k}{r}\frac{s}{k(s-k+r)}\\ = {} & \mathrm{II}_1+\mathrm{II}_2,\end{aligned}$$

where we split $(\nu+t+r)=(\nu+t+k)+(r-k)$.

$$\begin{aligned}\mathrm{II}_1 = {} & \frac{s}{s-k-t}\sum_{\nu=0}^{s-k-t}(-1)^{s-k-t-\nu}\binom{s-k-t}{s-k-t-\nu}\binom{\nu+t+k-1}{\nu+t}\\ & \times\sum_{r=0}^{k-1}(-1)^{k+r}\binom{k}{r}\binom{s-k+r-1}{r+t},\end{aligned}$$

where

$$\begin{aligned} & \sum_{r=0}^{k-1}(-1)^{k+r}\binom{k}{r}\binom{s-k+r-1}{r+t}\\ & = \sum_{r=0}^{k}(-1)^{k+r}\binom{k}{k-r}\binom{s-k+r-1}{r+t}-\binom{s-1}{t+k}\\ & = \sum_{\mu=t}^{k+t}(-1)^{k+t-\mu}\binom{k}{k+t-\mu}\binom{s-k-1-t-\mu}{\mu}-\binom{s-1}{t+k}\\ & = (k+t)\text{th coefficient of } (1-x)^k(1-x)^{k+1+t-s-1}-\binom{s-1}{t+k}\\ & = (-1)^{k+t}\binom{2k+t-s}{k+t}-\binom{s-1}{t+k},\end{aligned}$$

and where

$$\begin{aligned} & \sum_{\nu=0}^{s-k-t}(-1)^{s-k-t-\nu}\binom{s-k-t}{s-k-t-\nu}\binom{\nu+t+k}{\nu+t}\\ & = \sum_{\mu=t}^{s-k}(-1)^{s-k-\mu}\binom{s-k-t}{s-k-\mu}\binom{\mu+k}{\mu}\end{aligned}$$

$$
\begin{aligned}
&= (s-k)\text{th coefficient of } (1-x)^{s-k-t}(1-x)^{-k-1} \\
&= (-1)^{s-k}\binom{s-2k-t-1}{s-k}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathrm{II}_1 &= \frac{s}{s-k-t}(-1)^{s-k}\binom{s-2k-t-1}{s-k} \\
&\quad\times\left\{(-1)^{k+t}\binom{2k+t-s}{k+t}-\binom{s-1}{k+t}\right\} \\
&= (-1)^{s+t}\frac{s}{s-k-t}\binom{s-2k-t-1}{s-k}\binom{2k+t-s}{k+t} \\
&\quad -(-1)^{s-k}\binom{s-2k-t-1}{s-k}\binom{s}{k+t},
\end{aligned}
$$

so that, using the formulas above,

$$
\mathrm{I}+\mathrm{II}_1 = (-1)^{s+t}\frac{s}{s-k-t}\binom{s-2k-t-1}{s-k}\binom{2k+t-s}{k+t}.
$$

$$
\begin{aligned}
\mathrm{II}_2 &= \frac{s}{s-k-t}\sum_{\nu=0}^{s-k-t}(-1)^{s-k-t-\nu}\binom{s-k-t}{s-k-t-\nu}\binom{\nu+t+k-1}{\nu+t} \\
&\quad\times\sum_{r=0}^{k-1}(-1)^{r+k-1}\binom{k-1}{r}\binom{s-k+r-1}{r+t},
\end{aligned}
$$

where

$$
\begin{aligned}
&\sum_{r=0}^{k-1}(-1)^{r+k-1}\binom{k-1}{r}\binom{s-k+r-1}{r+t} \\
&\quad= \sum_{\mu=t}^{k+t-1}(-1)^{k+t-1-\mu}\binom{k-1}{k+t-1-\mu}\binom{s-k-1-t+\mu}{\mu} \\
&\quad= (k+t-1)\text{th coefficient of } (1-x)^{k-1}(1-x)^{k+t+1-s-1} \\
&\quad= (-1)^{k+t-1}\binom{2k+t-s-1}{k+t-1},
\end{aligned}
$$

and where

$$
\sum_{\nu=0}^{s-k-t}(-1)^{s-k-t-\nu}\binom{s-k-t}{s-k-t-\nu}\binom{\nu+t+k-1}{\nu+t}
$$

$$
\begin{aligned}
&= \sum_{\mu=t}^{s-k} (-1)^{s-k-\mu} \binom{s-k-t}{s-k-\mu} \binom{k-1+\mu}{\mu} \\
&= (s-k)\text{th coefficient of } (1-x)^{s-k-t}(1-x)^{-k} \\
&= (-1)^{s-k} \binom{s-2k-t}{s-k}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathrm{II}_2 &= \frac{s}{s-k-t}(-1)^{s+t-1} \binom{s-2k-t}{s-k} \binom{2k+t-s-1}{k+t-1} \\
&= -(-1)^{s+t} \frac{s}{s-k-t} \binom{s-2k-t-1}{s-k} \binom{2k+t-s}{k+t}.
\end{aligned}
$$

Altogether, using the formulas above,

$$
\alpha_k = \mathrm{I} + \mathrm{II}_1 + \mathrm{II}_2 = 0,
$$

and this completes the proof of Lemma 1. $\square$

Now, we can proceed with the proof of Theorem 1. Let $\mathscr{A} = (a_{\mu\nu})$ be given by (2). We define matrices $(\alpha_{t\mu}), (\beta_{\mu s}) \in \mathbb{R}^{n\times n}$ by

$$
\begin{aligned}
\alpha_{t\mu} &:= \begin{cases} \alpha_{t\mu}(x) = \displaystyle\sum_{\nu=0}^{\mu-t} \binom{\mu}{\nu} \binom{\mu}{\mu-t-\nu} x^{\nu} & \text{for } 0 \leqslant t \leqslant \mu \leqslant n, \\ 0 & \text{for } 0 \leqslant \mu < t \leqslant n, \end{cases} \\
\beta_{\mu s} &:= \begin{cases} \beta_{\mu s}(x) = f_\mu^s(x) & \text{for } 0 \leqslant \mu \leqslant s \leqslant n, \\ 0 & \text{for } 0 \leqslant s < \mu \leqslant n. \end{cases}
\end{aligned}
$$

Then, it follows from Eq. (6) of Lemma 1 that

$$
\sum_{\mu=0}^{n} \alpha_{t\mu}\beta_{\mu s} = \delta_{st} \quad \text{for } 0 \leqslant t, s \leqslant n, \tag{7}
$$

i.e., $(\alpha_{t\mu})$ is the inverse matrix of $(\beta_{\mu s})$ for $x \in \mathbb{R}$. Next, we introduce the generating functions

$$
\begin{aligned}
f_t(x) &:= \sum_{k\in\mathbb{Z}} r_t(k+t)x^k, \\
g_t(x) &:= \sum_{k\in\mathbb{Z}} (-1)^t a_{k+1,k+1+t} x^k \quad \text{for } 0 \leqslant t \leqslant n.
\end{aligned}
$$

Then, by Eq. (2),

$$\begin{aligned} g_t(x) &= \sum_{k\in\mathbb{Z}} \sum_{\mu=t}^{n} \sum_{\nu=t}^{\mu} \binom{\mu}{\nu}\binom{\mu}{\nu-t} r_\mu(k+\nu-\mu+\mu) x^{k+\nu-\mu} x^{\mu-\nu} \\ &= \sum_{\mu=t}^{n} \sum_{s=0}^{\mu-t} \binom{\mu}{\mu-s}\binom{\mu}{\mu-t-s} x^s f_\mu(x) \\ &= \sum_{\mu=0}^{n} \alpha_{t\mu}(x) f_\mu(x). \end{aligned}$$

Hence, by (7) and (5), for $0 \leqslant \mu \leqslant n$,

$$\begin{aligned} &\sum_{k\in\mathbb{Z}} r_\mu(k+\mu)x^k \\ &\quad = f_\mu(x) \\ &\quad = \sum_{s=0}^{n} \beta_{\mu s} g_s(x) \\ &\quad = \sum_{m\in\mathbb{Z}} \sum_{s=\mu}^{n} (-1)^{s-\mu} \left\{ \binom{s}{\mu} + \sum_{l=1}^{s-\mu} \frac{s}{l} \binom{\mu+l-1}{l-1}\binom{s-l-1}{s-\mu-l} x^l \right\} \\ &\qquad \times (-1)^s a_{m+1,m+1+s} x^m. \end{aligned}$$

By comparing the coefficient of x^k, we obtain formula (3), which completes the proof of Theorem 1. □

3. Corresponding eigenvalue problems and quadratic functionals

For a given difference operator of the form (1) we consider the following *discrete* Sturm–Liouville eigenvalue problem with eigenvalue parameter λ;

$$L(y)_k := \sum_{\mu=0}^{n} (-\Delta)^\mu \left\{ r_\mu(k)\Delta^\mu y_{k+1-\mu} \right\} = \lambda y_{k+1} \quad \text{for } 0 \leqslant k \leqslant N-n \tag{8}$$

with the boundary conditions

$$y_{1-n} = \cdots = y_0 = y_{N+2-n} = \cdots = y_{N+1} = 0, \tag{9}$$

where N and n are fixed integers with $1 \leqslant n \leqslant N$. Clearly, λ is an eigenvalue of (8) and (9), if there exists a corresponding eigenvector $y = (y_k)_{k=1}^{N+1-n} \in \mathbb{R}^{N+1-n}\backslash\{0\}$, such that (8) holds under condition (9). We need the following notation.

Notation. For $m > n \geqslant 1$, let $\mathscr{A}_m$ denote the submatrix of size $(m-n)\times(m-n)$ of the infinite matrix $\mathscr{A}$ from Section 2 in the left upper corner, i.e.,

$$\mathscr{A}_m = (a_{\mu\nu})_{\mu,\nu=1}^{m-n} \in \mathbb{R}^{(m-n)\times(m-n)},$$

where the matrix elements $a_{\mu\nu}$ are defined by (2) for $|\mu-\nu| \leqslant n$, and where $a_{\mu\nu} = 0$ for $|\mu-\nu| > n$.

Remark 1.

(i) Note that $\mathscr{A}_m$ is a banded, symmetric matrix of size $(m-n)\times(m-n)$ with bandwidth $2n+1$. Moreover, by Theorem 1, every matrix of this kind corresponds uniquely to a Sturm–Liouville operator, given by (1), via formulas (2) and (3) of Theorem 1. Assertion (i) of Theorem 1 states that the discrete Sturm–Liouville eigenvalue problem (8) and (9) is equivalent with the *algebraic eigenvalue problem*

$$\mathscr{A}_{N+1} y = \lambda y, \quad y \in \mathbb{R}^{N+1-n}. \tag{10}$$

(ii) If $N = 2n$, then *every* real, symmetric matrix $\mathscr{A} \in \mathbb{R}^{(n+1)\times(n+1)}$ has bandwidth $2n+1$. Therefore, by Theorem 1, every such matrix $\mathscr{A} = \mathscr{A}_{N+1}$ corresponds to a suitable Sturm–Liouville eigenvalue problem (8) and (9). Hence, all our results below apply to every real, symmetric matrix. But, of course, numerical algorithms resulting e.g. from Theorem 2 become efficient only if $N \gg n$, but not for "full" matrices, i.e., if $N \leqslant 2n$.

Lemma 2. *Let* $y \in \mathbb{R}^{N+1-n}$, *and suppose* (8) *and* (9). *Then*

$$y^{\mathrm{T}} \mathscr{A}_{N+1} y = \mathscr{F}(y),$$

where the quadratic functional $\mathscr{F}(y)$ *is defined by*

$$\mathscr{F}(y) := \sum_{k=0}^{N} \sum_{\mu=0}^{n} r_\mu(k) (\Delta^\mu y_{k+1-\mu})^2. \tag{11}$$

Proof. Using formula (4) for finite differences we obtain that

$$\begin{aligned}\mathscr{F}(y) &= \sum_{\kappa=0}^{n} \sum_{\mu=0}^{n} r_\mu(\kappa) \sum_{l=0}^{\mu} \binom{\mu}{l} (-1)^{\mu+l} y_{\kappa+1-\mu+l} \sum_{s=0}^{\mu} \binom{\mu}{s} (-1)^{\mu+s} y_{\kappa+1-\mu+s} \\ &= \sum_{r,\rho=1}^{N+1-n} \tilde{a}_{r\rho} y_r y_\rho\end{aligned}$$

with a symmetric matrix $(\tilde{a}_{\mu\nu}) \in \mathbb{R}^{(N+1-n)\times(N+1-n)}$, where, for $0 \leqslant k \leqslant N-n$ and $0 \leqslant t \leqslant n$ (observe (9) and put $k+1 = \kappa+1-\mu+l \geqslant 1,\ k+1+t = \kappa+1-\mu+s$, thus $t = s-l$; let $t \geqslant 0$; $\nu = \mu - l \geqslant 0$, so that $l = \mu-\nu \geqslant 0,\ \kappa = k+\nu,\ s = t+l = \mu-\nu+t \leqslant \mu,\ t \leqslant \nu \leqslant \mu$),

$$\tilde{a}_{k+1,k+1+t} = \sum_{\mu} \sum_{\kappa} r_\mu(\kappa)(-1)^{l+s} \binom{\mu}{l} \binom{\mu}{s}$$

$$= (-1)^t \sum_{\mu=t}^{n} \sum_{\nu=t}^{\mu} r_\mu(k+\nu) \binom{\mu}{\nu} \binom{\mu}{\nu-t},$$

and this equals $a_{k+1,k+1+t}$ by formula (2). This proves Lemma 2, because $\mathscr{A}$ is symmetric. □

4. The associated Hamiltonian system

In this section we formulate the connection between the discrete Sturm–Liouville equation (8) and an associated Hamiltonian difference system (cf. [3, Proposition 5] or [2, Lemma 2]), and we introduce the important notions of *conjoined bases* and *focal points* of it (cf. [3, Definitions 1 and 3], or [4, Definitions 1 and 3]). Moreover, the *multiplicity* of focal points is defined according to [7]. But it will turn out that these multiplicities always equal one for Hamiltonian systems, which we treat here, i.e., which originate from Sturm–Liouville equations.

The equivalence of Sturm–Liouville equations with Hamiltonian systems requires the assumption

$$r_n(k) \neq 0 \quad \text{for all } k, \tag{12}$$

which will be made in the sequel.

Lemma 3. *Assume* (12). *A vector* $y = (y_k)_{1-n}^{N+1} \in \mathbb{R}^{N+1-n}$ *solves the Sturm–Liouville difference equation* (8) (*for* $0 \leqslant k \leqslant N-n$) *if and only if* (x, u) *solves the Hamiltonian difference system*

$$\Delta x_k = A x_{k+1} + B_k u_k, \quad \Delta u_k = (C_k - \lambda \tilde{C}) x_{k+1} - A^{\mathrm{T}} u_k \tag{13}$$

for $0 \leqslant k \leqslant N$, *where we use the following notation*:

$A, B_k, C_k, \tilde{C}$ *are* $n \times n$*-matrices defined by*

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \end{pmatrix}, \quad B_k = \frac{1}{r_n(k)} B \ \textit{ with } B = \operatorname{diag}(0, \ldots, 0, 1),$$

$$C_k = \operatorname{diag}(r_0(k), \ldots, r_{n-1}(k)), \quad \textit{and} \quad \tilde{C} = \operatorname{diag}(1, 0, \ldots, 0),$$

for $0 \leqslant k \leqslant N$, *and* $x_k = (x_k^{(\nu)})_{\nu=0}^{n-1}$, $u_k = (u_k^{(\nu)})_{\nu=0}^{n-1} \in \mathbb{R}^n$ *are defined by*

$$x_k^{(\nu)} = \Delta^\nu y_{k-\nu}, \; u_k^{(\nu)} = \sum_{\mu=\nu+1}^{n} (-\Delta)^{\mu-\nu-1} \left\{ r_\mu(k) \Delta^\mu y_{k+1-\mu} \right\} \tag{14}$$

for $0 \leqslant \nu \leqslant n-1,\ 0 \leqslant k \leqslant N+1$ *with suitably chosen* $y_{N+2}, \ldots, y_{N+n+1}$ (*which are used for* $u_{N+2-n}, \ldots, u_{N+1}$).

Definition 1. Assume that (12) holds.

(i) A pair $(X, U) = (X_k, U_k)_{k=0}^{N+1}$ is called a conjoined basis of (13), if the real $n \times n$-matrices X_k, U_k solve (13) for $0 \leqslant k \leqslant N$, and if

$$X_0^{\mathrm{T}} U_0 = U_0^{\mathrm{T}} X_0 \quad \text{and} \quad \operatorname{rank}(X_0^{\mathrm{T}}, U_0^{\mathrm{T}}) = n \text{ holds.}$$

(ii) Suppose that (X, U) is a conjoined basis of (13) and let $0 \leqslant k \leqslant N$. We say that X has no focal point in the interval $(k, k+1]$ if

$$\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k \quad \text{and} \quad D_k := X_k X_{k+1}^{\dagger} \tilde{A} B_k \geqslant 0 \text{ holds,}$$

where $\tilde{A} := (I - A)^{-1}$. Moreover, if $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$ and $D_k \ngeqslant 0$ (i.e., the open interval $(k, k+1)$ contains a focal point of X), then $\operatorname{ind} D_k$ is called the multiplicity of "the" focal point of X in the interval $(k, k+1)$ (one could also say the number of focal points of X in $(k, k+1)$).

Remark 2.

(i) For a matrix M we denote by $\operatorname{Ker} M$ the kernel of M, $\operatorname{ind} M$ denotes the index of M, i.e., the number of negative eigenvalues of M, provided M is symmetric (and real), and $M^{\dagger}$ denotes the Moore–Penrose inverse of M (cf. [10, p. 421]). Moreover, $M \geqslant 0$ means that M is symmetric (and real) and non-negative definite, i.e., $\operatorname{ind} M = 0$. Observe that D_k is symmetric if $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$ (cf. [3, Proposition 1] or [4, Theorem 1]). Moreover, I denotes the identity matrix of suitable size.

(ii) For our Sturm–Liouville difference equations the multiplicity of focal points, which we defined only in case $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$, always equals 1, because $\operatorname{rank} D_k \leqslant \operatorname{rank} B = 1$.

5. The recursion formula

In this section we derive a recursion formula, which may be used to calculate numerically the determinant of symmetric, banded matrices. First we need some auxiliary results.

Lemma 4. *Suppose* (12), *and assume that* (X, U) *is the so-called principal solution of* (13), *i.e., the* $n \times n$*-matrices* $X_k = X_k(\lambda)$, $U_k = U_k(\lambda)$ *satisfy* (13) *with*

$$X_0 = 0, \quad U_0 = I. \tag{15}$$

Then the following assertions hold.

(i) $\operatorname{rank} X_k = k$, $D_k = 0$, $\operatorname{Ker} X_{k+1} \subset \operatorname{Ker} X_k$ *for* $k = 0, \ldots, n-1$.

(ii) $\det X_n = \{r_n(0) \cdots r_n(n-1)\}^{-1} \neq 0$.

(iii) $\det X_k(\lambda) \neq 0$ *for all* $n \leqslant k \leqslant N+1$, *and all* $\lambda \in \mathbb{R} \backslash \mathcal{N}$ *with a* finite *set* $\mathcal{N}$.

(iv)

$$D_k(\lambda) = \frac{1}{r_n(k)} \frac{\det X_k(\lambda)}{\det X_{k+1}(\lambda)} \quad \textit{for } n \leqslant k \leqslant N, \quad \lambda \in \mathbb{R},$$

provided $\det X_{k+1}(\lambda) \neq 0$, *where* $D_k(\lambda)$ *is defined as in Definition* 1, *i.e.*, $D_k(\lambda) = X_k(\lambda) X_{k+1}^{-1}(\lambda) \tilde{A} B_k$.

Proof. Assertion (i) is contained in [3, Proposition 6] or [5, Lemma 4].

For the proof of (ii) observe that, by Lemma 3 (note that Δ always operates with respect to k),

$$X_k = \left(\Delta^\mu y_{k-\mu}^{(\nu)}\right), \quad U_k = \left(\sum_{\rho=\mu+1}^{n} (-\Delta)^{\rho-\mu-1} \left\{r_\rho(k) \Delta^\rho y_{k+1-\rho}^{(\nu)}\right\}\right)$$

for $0 \leqslant \mu,\ \nu \leqslant n-1$, and that $X_0 = 0,\ U_0 = I$ by (15). This implies (use (4)) that, for $0 \leqslant \nu \leqslant n-1$,

$$y_0^{(\nu)} = y_{-1}^{(\nu)} = \cdots = y_{1-n}^{(\nu)} = 0,$$

and

$$\sum_{\rho=\nu+1}^{n} (-\Delta)^{\rho-\nu-1} \left\{r_\rho(k) \Delta^\rho y_{k+1-\rho}^{(\nu)}\right\} = 1,$$

$$\sum_{\rho=\mu+1}^{n} (-\Delta)^{\rho-\mu-1} \left\{r_\rho(k) \Delta^\rho y_{k+1-\rho}^{(\nu)}\right\} = 0,$$

if $k = 0,\ \nu < \mu \leqslant n-1$. Hence (note that $k = 0$), $r_n(0) \Delta^n y_{k+1-n}^{(\nu)} = 0$ for $0 \leqslant \nu < n-1$ (put $\mu = n-1$ above), so that $y_1^{(0)} = \cdots = y_1^{(n-2)} = 0$. Moreover, it follows inductively that

$$r_n(s) \Delta^{n+s} y_{k+1-n}^{(\nu)} = 0 \quad \text{for } 0 \leqslant s \leqslant n-2, \quad 0 \leqslant \nu \leqslant n-2-s,$$

so that $y_{s+1}^{(\nu)} = 0$ because $r_n(s) \neq 0$ by (13). Therefore,

$$(-1)^{n-\nu-1} r_n(n-\nu-1) \Delta^{2n-\nu-1} y_{k+1-n}^{(\nu)} = 1 \quad \text{for } 0 \leqslant \nu \leqslant n-1$$

so that

$$y_{n-\nu}^{(\nu)} = \frac{(-1)^{n-\nu-1}}{r_n(n-\nu-1)}.$$

Now, using suitable row operations, we obtain a matrix T with $\det T = 1$ such that

$$X_n T = \begin{pmatrix} y_n^{(0)} & \cdots & y_n^{(n-1)} \\ \vdots & & \vdots \\ (-1)^{n-1} y_1^{(0)} & \cdots & (-1)^{n-1} y_1^{(n-1)} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{(-1)^{n-1}}{r_n(n-1)} & \cdots & \star \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{(-1)^{n-1}}{r_n(0)} \end{pmatrix}.$$

Hence,

$$\begin{aligned} \det X_n = \det(X_n T) &= \frac{(-1)^{n-1}}{r_n(n-1)} \cdots \frac{(-1)^{n-1}}{r_n(0)} \\ &= \left\{r_n(0) \cdots r_n(n-1)\right\}^{-1} \neq 0. \end{aligned}$$

Next, let $y \in \mathbb{R}^{N+1-n} \setminus \{0\}$ and assume (9). Then, by Lemma 2,

$$\mathscr{F}(y) - \lambda \sum_{k=0}^{N+1} y_{k+1}^2 = y^{\mathrm{T}}(\mathscr{A}_{N+1} - \lambda I) y > 0,$$

provided λ is sufficiently small, more precisely: $\lambda < \min \sigma(\mathscr{A}_{N+1})$, where $\sigma(\mathscr{A}_{N+1})$ denotes the spectrum of $\mathscr{A}_{N+1}$. Therefore, by the *Reid Roundabout Theorem* (cf. [3, Satz 9] or [5, Theorem 2]),

$$\det X_{k+1}(\lambda) \neq 0 \quad \text{for } n \leqslant k \leqslant N \quad \text{and} \quad \lambda < \min \sigma(\mathscr{A}_{N+1}).$$

This implies assertion (iii), because the matrix elements of the matrix $X_k(\lambda)$ are poynomials in λ.

Finally, assertion (iv) is shown in [6, Lemma 4.1]. □

Lemma 5 (Picone's identity). *Suppose* (12), *let* (X, U) *be the principal solution of* (13), *i.e.*, (15) *holds, and assume that* $\det X_{k+1}(\lambda) \neq 0$ *for* $n \leqslant k \leqslant N$. *Then, for* $y \in \mathbb{R}^{N+1-n}$ *and under assumption* (9), *we have the formula* (*cf. the notation of Section* 3)

$$y^{\mathrm{T}}(\mathscr{A}_{N+1} - \lambda I) y = \sum_{k=n}^{N} r_n(k) \frac{\det X_{k+1}(\lambda)}{\det X_k(\lambda)} w_{k+1-n}^2, \tag{16}$$

where $w_\nu = y_\nu + \sum_{\mu=\nu+1}^{\nu+n} \alpha_\mu y_\mu$ *with suitable coefficients* $\alpha_\mu = \alpha_\mu(\nu, \lambda)$. *Hence,* $w = Ty$ *with*

$$T = \begin{pmatrix} 1 & \star & \star \\ \vdots & \ddots & \star \\ 0 & \cdots & 1 \end{pmatrix},$$

so that $\det T = 1$.

Proof. First, it follows from Lemma 2, Lemma 4(i), (ii), and [3, Proposition 1] or [4, Theorem 1] that

$$y^{\mathrm{T}}(\mathscr{A}_{N+1}-\lambda I)y=\mathscr{F}(y)-\lambda\sum_{k=n}^{N-n}y_{k+1}^2=\sum_{k=n}^{N}z_k^{\mathrm{T}}D_kz_k,$$

where $z_k=u_k-U_k(\lambda)X_k^{-1}(\lambda)x_k$ with x_k,u_k as in Lemma 3.

Therefore, by Lemma 4(iii),

$$y^{\mathrm{T}}(\mathscr{A}_{N+1}-\lambda I)y=\sum_{k=n}^{N}\frac{1}{r_n(k)}\frac{\det X_k(\lambda)}{\det X_{k+1}(\lambda)}z_k^{\mathrm{T}}Bz_k.$$

Let

$$Q_k:=U_k(\lambda)X_k^{-1}(\lambda)=(q_{\mu\nu}(k))_{\mu,\nu=0}^{n-1},\quad z_k=(z_k^{(\nu)})_{\nu=0}^{n-1}.$$

Then

$$z_k^{\mathrm{T}}Bz_k=\left(z_k^{(n-1)}\right)^2=\left\{u_k^{(n-1)}-\sum_{\nu=0}^{n-1}q_{n-1,\nu}(k)\Delta^\nu y_{k-\nu}\right\}^2$$

and

$$u_k^{(n-1)}=r_n(k)\Delta^n y_{k+1-n}=r_n(k)\left\{y_{k+1}+\cdots+(-1)^n y_{k+1-n}\right\}$$

by the notation of Lemma 3.

By (13), we have that $X_{k+1}=\tilde{A}(X_k+B_kU_k)$, and therefore

$$\begin{aligned}\det X_{k+1}&=\det(X_k+B_kU_k)\\&=\det\left(I+\frac{1}{r_n(k)}BQ_k\right)\det X_k\\&=\det\begin{pmatrix}1&\cdots&0\\\vdots&\ddots&\vdots\\\star&\cdots&1+q_{n-1,n-1}(k)/r_n(k)\end{pmatrix}\det X_k\\&=\frac{\det X_k}{r_n(k)}\left\{r_n(k)+q_{n-1,n-1}(k)\right\}.\end{aligned}$$

Hence,

$$\begin{aligned}z_k^{(n-1)}&=u_k^{(n-1)}-\sum_{\nu=0}^{n-2}q_{n-1,\nu}(k)\Delta^\nu y_{k-\nu}-\left\{\frac{\det X_{k+1}}{\det X_k}-1\right\}r_n(k)\Delta^{n-1}y_{k+1-n}\\&=(-1)^n r_n(k)\frac{\det X_{k+1}}{\det X_k}y_{k+1-n}+\sum_{\mu=k+2-n}^{k+1}\tilde{\alpha}_\mu y_\mu\end{aligned}$$

with suitable coefficients $\tilde{\alpha}_\mu=\tilde{\alpha}_\mu(k,\lambda)$. It follows that

$$\frac{1}{r_n(k)}\frac{\det X_k(\lambda)}{\det X_{k+1}(\lambda)}z_k^{\mathrm{T}}Bz_k=r_n(k)\frac{\det X_{k+1}(\lambda)}{\det X_k(\lambda)}\left\{y_{k+1-n}+\sum_{k+2-n}^{k+1}\alpha_\mu y_\mu\right\}^2$$

with certain α_μ, and this completes the proof of Lemma 5. □

Now, we can prove the main result of this section.

Theorem 2 (Recursion). *Assume* (12), *and let* (X, U) *be the principal solution of* (13), *i.e.*, (15) *holds.*

Then, *with the notation of Sections* 3 *and* 4,

$$\det(\mathscr{A}_{N+1} - \lambda I) = r_n(0) \cdots r_n(N) \det X_{N+1}(\lambda) \quad \textit{for all } \lambda \in \mathbb{R}, \tag{17}$$

where, *by* (13) *and* (15), $X_{N+1}(\lambda)$ *is given by the recursion*

$$X_{k+1} = \tilde{A}(X_k + B_k U_k), \quad U_{k+1} = (C_k - \lambda \tilde{C}_k) X_{k+1} + (I - A^{\mathrm{T}}) U_k$$

for $0 \leqslant k \leqslant N$ *with* $X_0 = 0, U_0 = I$.

Proof. Let $\lambda \in \mathbb{R} \backslash \mathscr{N}$ according to assertion (iii) of Lemma 4. Then, by Lemma 5 and assertion (ii) of Lemma 4,

$$\begin{aligned}\det(\mathscr{A}_{N+1} - \lambda I) &= r_n(n) \frac{\det X_{n+1}(\lambda)}{\det X_n(\lambda)} \cdots r_n(N) \frac{\det X_{N+1}(\lambda)}{\det X_N(\lambda)} \\ &= r_n(n) \cdots r_n(N) \frac{\det X_{N+1}(\lambda)}{\det X_n} \\ &= r_n(0) \cdots r_n(N) \det X_{N+1}(\lambda).\end{aligned}$$

Hence, (17) holds for all $\lambda \in \mathbb{R} \backslash \mathscr{N}$, and by continuity it is true for all $\lambda \in \mathbb{R}$. □

Remark 3. Assume (12), let (X, U) be the principal solution of (13), and using the notation of Section 3 define

$$f_k(x) := \det(\mathscr{A}_k - xI) \quad \text{for } n+1 \leqslant k \leqslant N+1 \quad \text{and} \quad f_n(x) \equiv 1.$$

Let $\lambda \in \mathbb{R}$ such that $f_k(\lambda) \neq 0$ for $n+1 \leqslant k \leqslant N+1$. Then, by the previous theorem, for $n \leqslant k \leqslant N$,

$$f_{k+1}(\lambda) = r_n(0) \cdots r_n(k) \det X_{k+1}(\lambda)$$

so that

$$\frac{f_{k+1}(\lambda)}{f_k(\lambda)} = r_n(k) \frac{\det X_{k+1}(\lambda)}{\det X_k(\lambda)}.$$

Then, Picone's identity (16) reads as follows:

$$y^{\mathrm{T}}(\mathscr{A}_{N+1} - \lambda I) y = \sum_{k=n}^{N} \frac{f_{k+1}(\lambda)}{f_k(\lambda)} w_{k+1-n}^2, \tag{18}$$

where $w_\nu = y_\nu + \sum_{\mu=\nu+1}^{\nu+n} \alpha_\mu y_\mu$ with suitable coefficients $\alpha_\mu = \alpha_\mu(\nu, \lambda)$. This formula (18) is the well-known formula of Jacobi [8, Chapter X, Section 3, formulae (20), (25), and (26)]. Hence, the general Picone's identity for linear Hamiltonian difference systems of Bohner (cf. [3, Proposition 1] or [4, Theorem 1]) is just a generalization of the old formula of Jacobi.

6. The oscillation result

First, we formulate a known result [8, Chapter X Section 3, Theorem 2 (Jacobi)], which is a direct consequence of Jacobi's formula (18) and Sylvester's law of inertia.

Lemma 6. *Assume* (12), *and with the notation of Section* 3 *define as in Remark* 3 *polynomials* $f_k(x)$ *by*

$$f_k(x) := \det(\mathscr{A}_k - xI) \quad \textit{for } n+1 \leqslant k \leqslant N+1 \quad \textit{and} \quad f_n(x) \equiv 1. \tag{19}$$

Let $\lambda \in \mathbb{R}$ *such that* $f_k(\lambda) \neq 0$ *for* $n+1 \leqslant k \leqslant N+1$.

Then the number of zeros of $f_{N+1}(x)$ (*including multiplicities*), *which are less than* λ, *equals the number of sign changes of* $\{f_k(\lambda)\}$ *for* $n \leqslant k \leqslant N+1$, *i.e.*, $\{f_k(\lambda)\}$ *is a* "*Sturmian chain*".

Remark 4. Another "formulation" of this statement reads as follows (cf. [8, Chapter X, Section 3, Theorem 2]). Let $\mathscr{B} = (b_{\mu\nu}) \in \mathbb{R}^{m\times m}$ be any real symmetric matrix and define

$$d_k = \det \begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{kk} \end{pmatrix} \quad \text{for } k = 1, \ldots, m.$$

If $d_k \neq 0$ for $1 \leqslant k \leqslant m$, then the number of negative eigenvalues of $\mathscr{B}$, i.e., ind $\mathscr{B}$, equals the number of sign changes in the finite sequence $\{1, d_1, \ldots, d_m\}$. Note also that Lemma 6 yields the Poincaré separation theorem (cf. [9, Theorem 8.5-1]; [10, Corollary 4.3.16]; or [13, Sätze 4.9 and 4.11]).

For the numerical treatment of the eigenvalue problem for real, symmetric tridiagonal matrices (i.e., our case $n = 1$) one uses the above results (cf. [9, Chapter 8] or [14, p. 299]). Actually, Wilkinson [14, p. 300] cites

> *The strict separation of the zeros for symmetric tridiagonal matrices with non-zero super-diagonal elements forms the basis of one of the most effective methods of determining the eigenvalues* (*Givens*, 1954).

In view of our results here the corresponding algorithm may now be applied to any symmetric banded matrix without reduction to tridiagonal form.

In the sequel we consider the eigenvalue problem (8) and (9) of Section 3 for some fixed $\lambda \in \mathbb{R}$. It follows immediately from assertion (i) of Theorem 1 and from Lemma 3 that the eigenvalue problem is equivalent with the eigenvalue problem for the corresponding Hamiltonian difference system (13) with the boundary conditions $x_0 = x_{N+1} = 0$, i.e.,

$$\text{(E)} \qquad \Delta x_k = Ax_{k+1} + B_k u_k,\ \ \Delta u_k = (C_k - \lambda\tilde{C})x_{k+1} - A^{\mathrm{T}} u_k \\ \text{for } 0 \leqslant k \leqslant N, \text{ with } x_0 = x_{N+1} = 0.$$

If (X, U) is the principal solution of (13), then every solution (x, u) of (13) with $x_0 = 0$ satisfies

$$x_k = X_k(\lambda)u_0, \quad u_k = U_k(\lambda)u_0 \quad \text{for } 0 \leqslant k \leqslant N + 1.$$

Therefore, we have the following lemma.

Lemma 7. *Assume* (12), *and suppose that* (X, U) *is the principal solution of* (13), *i.e.,* (15) *is satisfied. Then the following assertions hold*:

(i) *A number* $\lambda \in \mathbb{R}$ *is an eigenvalue of the eigenvalue problem* (8) *and* (9) *if and only if*

$$\det X_{N+1}(\lambda) = 0,$$

and the dimension of the kernel of $X_{N+1}(\lambda)$ *is the multiplicity of* λ.

(ii) *A vector* $y = (y_k)_{k=1}^{N+1-n} \in \mathbb{R}^{N+1-n}\backslash\{0\}$ *is an eigenvector of an eigenvalue* $\lambda \in \mathbb{R}$ *of* (8) *and* (9) *if and only if* $x_k = X_k(\lambda)c,\ u_k = U_k(\lambda)c$ *for* $0 \leqslant k \leqslant N + 1$ *and some* $c \in \operatorname{Ker} X_{N+1}(\lambda)\backslash\{0\}$, *where* x_k *and* u_k *are defined by* (14) *and where* (9) *is assumed to be satisfied.*

This lemma shows that the algebraic eigenvalue problem (10) for symmetric, banded matrices can be treated completely via the corresponding eigenvalue problem (8) and (9) for Sturm–Liouville difference equations (using Theorem 1) and via the recursion of Theorem 2 (using Lemma 7).

Let us now formulate the main result of this section, which is in view of the above theory essentially a reformulation of Lemma 6 in terms of the oscillation of the principal solution of (13). This result, Theorem 3 below, has a complete analog for (continuous) Hamiltonian systems (cf. [11, Theorem 7.2.2 “global oscillation theorem” for the principal solution at 0]). But before we make the following remark.

Remark 5. By (i) of Lemma 7, the multiplicity of an eigenvalue λ is given by the dimension of the kernel of $X_{N+1}(\lambda)$. Hence, it is an integer in $\{1, \ldots, n\}$. Moreover, by (ii) of Remark 2, the focal points of $X(\lambda)$ are all simple, i.e., of multiplicity 1, and, by Lemma 4(iv) and Definition 1(ii), their number in $(0, N + 1]$ equals the number of elements of the set

$$\left\{k : n \leqslant k \leqslant N \text{ with } r_n(k)\frac{\det X_{k+1}(\lambda)}{\det X_k(\lambda)} < 0\right\},$$

provided that $\det X_{k+1}(\lambda) \neq 0$ for $n \leqslant k \leqslant N$.

Combining Theorem 2, Lemma 6 and Remarks 3 and 5, we obtain the main result of this section.

Theorem 3 (Oscillation). *Assume* (12), *and let* (X, U) *be the principal solution of* (13) (*i.e.*, (15) *holds*), *where we use the notation of Lemma* 3. *Moreover*, *let* $\lambda \in \mathbb{R}$ *be such that* $\det X_{k+1}(\lambda) \neq 0$ *for* $n \leqslant k \leqslant N$, *i.e.*, $\lambda \in \mathbb{R} \setminus \mathscr{N}$ *as in assertion* (iii) *of Lemma* 4.

Then, *the number of eigenvalues* (*including multiplicities*) *of the eigenvalue problem* (8), (9) (*or* (10)) *of Section* 3, *which are less than* λ, *equals the number of focal points of* $X(\lambda)$ *in the interval* $(0, N+1]$.

Finally, let us mention that this statement is a special case of the general oscillation theorem for Hamiltonian difference systems from [7]. More precisely, it is the special case, where the Hamiltonian system corresponds to a Sturm–Liouville difference equation via Lemma 3.

References

[1] L. Berg, Lineare Gleichungssysteme mit Bandstruktur, Carl Hanser, München, 1986.
[2] M. Bohner, Controllability and disconjugacy for linear Hamiltonian difference systems, in: Proceedings of the First International Conference on Difference Equations, San Antonio, 1994, pp. 65–77.
[3] M. Bohner, Zur Positivität diskreter quadratischer Funktionale, Dissertation, Ulm, 1995.
[4] M. Bohner, Linear Hamiltonian difference systems: disconjugacy and Jacobi-type conditions, J. Math. Anal. Appl. 199 (1996) 804–826.
[5] M. Bohner, On disconjugacy for Sturm–Liouville difference equations, J. Differential Equations Appl. 2 (1996) 227–237.
[6] M. Bohner, O. Došlý, Positivity of blockridiagonal matrices, SIAM J. Matrix Anal. Appl. 20 (1999) 182–195.
[7] M. Bohner, O. Došlý, W. Kratz, An oscillation theorem for discrete eigenvalue problems, to appear.
[8] F.R. Gantmacher, The Theory of Matrices, vol. 1, Chelsea, New York, 1959 (originally published in Moscow, 1954).
[9] G.H. Golub, C.F. Van Loan, Matrix Computations, John Hopkins University Press, Baltimore, MD, 1983.
[10] R.A. Horn, C.A. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[11] W. Kratz, Quadratic Functionals in Variational Analysis and Control Theory, Akademie, Berlin, 1995.
[12] W.T. Reid, Sturmian Theory for Ordinary Differential Equations, Springer, New York, 1980.
[13] H.R. Schwarz, H. Rutishauser, E. Stiefel, Numerik symmetrischer Matrizen, Teubner, Stuttgart, 1972.
[14] J.H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, UK, 1965.